JP 1998-232934A

(11)Publication number: 10-232934
(43)Date of publication of application: 02.09.1998
(21)Application number: 09-033704
(22)Date of filing: 18.02.1997
(71)Applicant: TOSHIBA CORP
(72)Inventor: FUKUI KAZUHIRO , YAMAGUCHI OSAMU
(54) Name of this invention: **FACE IMAGE REGISTERING DEVICE AND ITS METHOD**

(57)Abstract:
PROBLEM TO BE SOLVED: To provide the face image registering device which can exclude the registration of an unnecessary face image without placing any burden on a user and also increase dictionary precision.
SOLUTION: An information terminal device 10 with security consists of an image input part 11, a face area extraction part 12, a state recognition part 13, a recognition part 14, a dictionary generation part 15, an access control part 16, dictionary images 17, and a specific operation detection part 18. After the specific operation detection part 18 detects specific operation of the user, the face area extraction part 12 extracts an image of a face area from an image inputted from the image input part 11 and a dictionary generation part 15 judges whether or not the image is valuable enough to register as a dictionary image 17.

[Detailed Description of the Invention]
[0001]
[Field of the Invention] This invention relates to the face image registration equipment which extracts a face field out of the inputted image, and carries out dictionary registration of the face automatically, or its approach.
[0002]
[Description of the Prior Art] A face image recognition technique is a component engineering indispensable to construction of a human interface or a security system. Face image recognition consists of a face field extract, a face focus extract, and discernment. About the latest technical trend, reference [latest research trend [of a Shiono **** Sanada / Hidehiko /" personal authentication technique]", Shingaku Giho OSF92-17.] is detailed. The conventional face image identifying method can be classified into two of the roughly divided followings.
[0003] One is the approach of calculating similarity with a candidate person's feature

vector which parameterizes the location of the focus, such as an eye, a nose, and opening, a configuration, and size, generates a feature vector, and is registered beforehand. The person showing a dictionary vector with the highest similarity is discriminated from a person in question. These are classified into the structural-analysis-technique.

[0004] Another approach is an approach based on the similarity of the pattern of the image by which a location and size were normalized by geometrical conversion of the two-dimensional affine conversion on the basis of the focus, such as a pupil and a nose, etc., and the dictionary normalization image registered beforehand. The person who expresses a dictionary image with the highest similarity like the former is discriminated from a person in question. These can be classified into the technique like a pattern according to the approach put in practical use in the conventional character recognition.

[0005] In both [of the two above-mentioned technique] cases, statistical distance, such as a correlation value between images and Euclidean distance in the inside of a feature space, is used as similarity. Approaches, such as various pattern recognition theory put in practical use in character recognition, for example, a subspace method, [the Hidemitsu Ogawa written by ERUKKI OYA and Sato **** and "pattern recognition and subspace-method" Sangyo Tosho Publishing (1986)], and compound similarity, are applicable to this discernment processing. The configuration of the identifying method is detailed to [Funakubo ** "pattern recognition" KYORITSU SHUPPAN (1991)], [Taizo Iijima "pattern recognition theoretical" Morikita Shuppan (1989) etc.], etc.

[0006] In addition to the above-mentioned face recognition processing, on the occasion of actual application of the above-mentioned face recognition, registration of a user, i.e., a registrant's face dictionary generation, becomes indispensable. For example, in the handwriting recognition of Chinese character using a subspace method, the dictionary was generated from the study sample image of hundreds of sheets for every alphabetic character. By this invention, compared with an alphabetic character, fluctuation of configurations, such as lighting conditions, face sense, and expression change, or brightness is large, and, as for a three-dimension body like the target face, the study sample image of a large quantity is further needed [Murase **, sherry NAIYA, "the three-dimension object recognition by two-dimensional collating", IEICE TRANSACTIONS (D-II) J77-D-II, 11, pp.2179-2187, and 1994.].

[0007] The conventional dictionary generation Reference [red Matsushige, Sasaki **, Akio Fukamachi, Yasuhito Suenaga, "the robust method of identifying a transverse-plane face by shade image matching", IEICE TRANSACTIONS (D-II), J76-DII, 7, pp.1363-1373, 1993.], [Yoshie Komatsu, Yasuo Ariki, "logging recognition of

the face by the sense using a subspace method", PRU 95-191, pp.7-14, 1996.], [-- M.Turk and A.P.Pentland : [ ] -- "Face recognition using eigenfaces" -- Proc.CVPR 11, pp.453-458, and 1993. --] -- [Alex Pentland and Baback Moghaddam and Thad Starner, So that "View-based and modular eigenspaces for face recognition", CVPR'94, pp.84-91, 1994.], etc. may see The help was performing logging of a face and selection of an image from the image which turned and photoed a registrant's face under good lighting conditions in the direction of predetermined, for example, a transverse plane, 15 degrees of right and left, and 15 degree ** of vertical.

[0008] A physiological change of the effect of the condition of the day, expression change and hair, and a mustache etc. tends to produce a face. Therefore, in order to realize a high recognition rate, it is necessary to update a face dictionary for every fixed period. However, the frequent renewal of a dictionary will raise a user's burden. So, in order to make face recognition easy to use, it becomes important how a burden is mitigated and registered without making a user conscious.

[0009] For example, when it is interrupted temporarily and leaves the activity in an information terminal, the function to prevent others' access is considered. although using the dictionary generated beforehand is also considered -- a principal -- it is more effective to use the dictionary image generated as much as possible immediately before, in order to gather identification and the rate of others abatement. However, in order to make a user's burden high, dictionary registration of carrying out dictionary registration at every **** is carried out without a user's making it conscious while accessing information machines and equipment.

[0010] However, the study sample collected without making a registrant conscious contains the data which are not effective as study data the case where an eye is usually closed, when expression is changing a lot. If dictionary generation is performed from the study sample in which these error images are contained, the precision of a dictionary will fall. In order to raise dictionary precision, it is necessary to increase the number of study data, and real-time dictionary generation cannot be coped with in selection by the help.

[0011]

[Problem(s) to be Solved by the Invention] Conventionally, the help was performing generation of a face dictionary using predetermined, for example, a transverse plane, 15 degrees of right and left, and the face image that was made to turn a registrant's face in the direction for 15 degrees of every upper and lower sides, and was photoed in advance under good lighting conditions. A face changes with time amount. Therefore, in order to realize a high recognition rate, it is necessary to update a face dictionary for every fixed

period. However, performing frequent renewal of a dictionary will raise a user's burden.

[0012] About the automatic study sample collection for dictionary generation, if an approach [finishing / application / already] [Japanese Patent Application No. No. 61463 [eight to]] is applied, it is realizable. However, the technical problem that much data which are not effective in dictionary generation are contained remains in the study data collected without still making a registrant-ed conscious. In order to raise dictionary precision, it is necessary to increase the number of study data, and real-time dictionary generation is difficult to realize in selection by the help.

[0013] Then, this invention provides a user with a burden, ****, the face image registration equipment that can eliminate registration of an unnecessary face image and can raise dictionary precision, and its approach.

[0014]

[Means for Solving the Problem] In the face image registration equipment which has a dictionary means by which this invention has registered a person's face image A face field extract means to extract the input face image which is an image of said person's face field out of the image inputted by image input means to input said person's image, and said image input means, A decision means to judge whether the input face image extracted by said face field extract means is a face image which can be registered into said dictionary means, It is face image registration equipment characterized by consisting of a registration means made to register into said dictionary means with said decision means by using as said face image the input face image judged that registration is possible.

[0015] Moreover, this invention is set to the face image registration approach for registering a person's face image. The face field extract step which extracts the input face image which is an image of said person's face field out of the image inputted in the image input step which inputs said person's image, and said image input step, The decision step the input face image extracted in said face field extract step judges it to be whether it is the face image in which said registration is possible, It is the face image registration approach characterized by consisting of a registration step into which the input face image judged that registration is possible in said decision step is made to register as said face image.

[0016] It judges whether it is effective as study data by comparing the criteria beforehand set up in the characteristic quantity which can be found from the face field extracted, for example as it is the above-mentioned invention. The image which normalized as characteristic quantity on the basis of the magnitude of the face focus, such as a pupil, a nostril, and ****, the amounts which can be found from positional

information, or these points is used. It becomes possible to remove an open beam image etc. automatically from a study image about the image from which expression changed a lot by this, the image which closed the eye, and opening.

[0017] Furthermore predetermined actuation of mouse actuation, keyboard grabbing, button grabbing, etc. is detected, and dictionary generation is performed only for the image inputted during actuation. It decreases [the time amount of useless study data collection] for being aimed only at under actuation and is efficient. Moreover, possibility of collecting the images which are not effective as dictionaries in the case of having turned to width etc. also decreases. According to this invention, the equipment and the approach of extracting a face field out of the inputted image, and carrying out dictionary registration automatically are realizable.

[0018]

[Embodiment of the Invention] Hereafter, one example of this invention is explained based on drawing 7 from drawing 1 .

[0019] In information terminal units, such as a personal computer, in case under an activity is interrupted temporarily and a terminal is left, it is a required function to prevent others' access. Therefore, this example explains taking the case of the information-machines-and-equipment terminal unit 10 which carried the function to prevent others' access at the time of ****.

[0020] Drawing 1 is the block diagram showing the outline of the information terminal unit 10 with security concerning this example.

[0021] This equipment 10 consists of the image input section 11, the face field extract section 12, the situation-recognition section 13, a private seal 114, the dictionary generation section 15, the access-control section 16, a dictionary image 17, and a predetermined actuation detecting element 18.

[0022] (Image input section 11) The image input section 11 is for inputting the image of the person who becomes an object for recognition, for example, consists of a TV camera. The image 01 inputted from this image input section 11 is digitized by the A/D converter, and is sent to the face field extract section 12. For example, a TV camera is installed in the lower part of the monitor of the information terminal unit 10. Or you may install in the rectangular head of a monitor.

[0023] (Face field extract section 12) The face field extract section 12 always continues extracting the face field image 02 from the input image sent from the image input section 11. while moving the standard face image (template) registered beforehand over a full screen in this example -- a correlation value -- calculating -- the highest correlation value -- **** -- let the field which is be a face field. Let the partial maximum

point of a correlation value be a face field candidate. Suppose that a face does not exist, when lower than the threshold to which the correlation value in this face field candidate was set. It is set as 30, specifically using a maximum correlation value as 100 as a threshold. Since it corresponds to sense change of a face, if two or more templates are used by compound similarity etc., a face field can be further extracted to stability. This processing may be transposed to the extraction method based on the color information described previously.

[0024] (Situation-recognition section 13) In the situation-recognition section 13, a user's condition (under an activity, ****, taking a seat) is discriminated from a time change of a face field extract result, and the change in dictionary generation mode and recognition mode is performed.

[0025] It is set as recognition mode at the periods from T0 to T1 the event of a face field beginning to be detected as shown in drawing 2 . the recognition processing described in this mode later -- a principal -- identification processing is performed. It becomes accessible when identified a principal at the T1 event.

[0026] When it changes to dictionary generation mode after T1 event and the dictionary generation signal from the predetermined actuation detecting element 18 is inputted, collection of study data is performed. It judges that it ****(ed) when it became impossible to have detected a fixed period face field, and a screen lock is started immediately. Or a screen lock may be started, after conveying the main point to a user with light, voice, etc. and seeing a user's reaction. Time of day T1 can be changed according to an object model or the class of access actuation here.

[0027] (Recognition section 14) The recognition section 14 consists of face focus extract section 14a, normalization image generation section 14b, and pattern-matching section 14c. The block diagram of the recognition section 14 is shown in drawing 3 .

[0028] In face focus extract section 14a, the focus, such as a pupil, a nose, and ****, is extracted from the inside of the extracted face field. The approach [Japanese Patent Application No. No. 61463 [eight to]] which combined the configuration information for which it has already applied, and pattern information is applicable.

[0029] The fundamental idea of this approach asks for the candidate of the focus using the configuration information that location precision is high, and verifies it by pattern matching. Since this approach positions using configuration information, it can expect a high location precision. Moreover, since matching which used the multi-template for selection of the right focus from a candidate group is applied, it is robust to fluctuation of the configuration brightness of the focus. Since pattern matching is carried out about processing speed only to the candidate who narrowed down with the

degree-of-separation filter with little count cost, compared with the approach of carrying out pattern matching of the whole, the drastic cutback of computational complexity is realizable.

[0030] In addition to this Approach [A.L.Yuille based on edge information, "Feature extractionfrom faces using deformable templates" and IJCV, vol.8 : 2 and pp.99-111, 1992.][Shizuo Sakamoto, the Miyao proton, Joji Tajima, "a focus extract of the eye from a face image", IEICE TRANSACTIONS D-II, Vol.J76-D-II, No.8, pp.1796-1804, August, 1993.], and a proper space method Applied Eigen feature Method [Alex Pentland, Baback Moghaddam, Thad Starner, and "View-based and modular eigenspaces for face recognition", The approach [the Sasaki **** Akamatsu **, Yasuhito Suenaga, "the method of alignment of the face using the color information for face image recognition", IE 91-2, pp.9-15, and 1991.] based on CVPR'94, pp.84-91, 1994.], and color information is applicable.

[0031] In normalization image generation section 14b, it normalizes on the basis of the focus. The example of the normalization processing on the basis of a pupil and a nostril is shown in drawing 4 . The sense of vector E1E2 is amended to parallel, further, from on c1 and c2, the main coordinate of a normalization image and breadth are set up by 2 twice vector E1E2, and a dip sets up the point CP of 1/3 by 2 twice vector c1c2.

[0032] In pattern-matching section 14c, it asks for pattern similarity as compared with a normalization image and the face image currently stored in the dictionary image. It identifies being a principal, when pattern similarity is higher than a reference value. In being smaller than criteria, it considers as others. As the approach of pattern matching, a correlation technique, a subspace method, a compound similarity method, etc. are applicable.

[0033] (Predetermined actuation detecting element 18) The predetermined actuation detecting element 18 detects whether the user is performing predetermined actuation of a keyboard, a carbon button push, or actuation of a mouse. The dictionary generation section 15 is controlled based on a detection result.

[0034] Since it is suitable to specifically carry out dictionary registration of the user performing predetermined actuation of keyboard grabbing, a carbon button push, actuation of a mouse, etc. since possibility that the transverse-plane image of a user's face is reflected to the image input means 11 is high when it detects, while performing predetermined actuation, a dictionary generation signal is transmitted to the dictionary generation section 16.

[0035] (Dictionary generation section 16) The dictionary generation section 16 consists of face focus extract section 14a, frame assessment section 16a, normalization image

generation section 14b, study image storage section 16b, and principal-component-analysis section 16c. With the recognition section 14, face focus extract section 14a and normalization image generation section 14b are common, and are used here. The block diagram of the dictionary generation section 16 is shown in drawing 5 .

[0036] If a dictionary generation signal is inputted from the predetermined actuation detecting element 18, as for dictionary generation, collection of study data will be performed.

[0037] In frame assessment section 16b, it judges whether the face field image 02 extracted from the information on the extracted face focus is effective in dictionary generation. Although three valuation bases are applied here, you may apply in individual or combination.

[0038] (i) The physical relationship of each focus is specified beforehand, and when this physical relationship collapses, it is judged that it is not effective as study data. For example, they are both pupils, a nostril, the physical relationship of ****, etc.

[0039] (ii) The proper face image and pattern similarity which were stored in the dictionary image 17 are made assessment. When similarity is smaller than criteria, it is judged that it is not effective. The homogeneity of study data will be raised, so that criteria are made high. When too high, it becomes impossible however, to correspond to change of the face sense etc.

[0040] Or the dictionary image generated from the study image normalized by the right focus in advance instead of the proper face image stored in the dictionary image 17 may be used.

[0041] If it furthermore takes into consideration to the face sense, it is possible to remove the face which turned to width extremely. In this case, when the comparison with the face dictionary pattern currently beforehand prepared for every face sense of a direction 1 - N is performed and it is in agreement with the dictionary of the predetermined face sense, it is judged that it is effective in dictionary generation.

[0042] (iii) The study data which investigated the relation of the focus extracted for every frame, and closed the eye are detected. The coordinate of the focus extracted newly can be expressed by the linear combination of the coordinate of the focus already extracted with the frames 1-4 before that. Therefore, if the focus is correctly extracted with all frames, the error at the time of carrying out linear combination will become small. Conversely, an error becomes large, when an eye is closed and eyebrows and a pupil are mistaken.

[0043] Drawing 6 is used and explained to a detail about this.

[0044] If a orthogonal projection model is assumed, the two-dimensional coordinate of the face focus seen from the direction of arbitration, without having a three-dimension model By the linear combination of the two-dimensional coordinate of the focus matched on the image of four sheets as shown in a formula (1) and (2) [S.Ullman which can be expressed, and R. Basri: "Recognition by Linear Combinations of Models, IEEE Trans.PAMI, Vol.13, No.10, pp.992-1006, 1991.][Yasuhiro Mukogawa, Yuichi Nakamura, Tomoichi Ota, "generation of the face image of the direction of arbitration using two photographs of his face", an information processing theory, Vol.37, and No. -- 4, pp.635-644, and 1996. --] . The two-dimensional coordinate value (X, Y) of the focus newly detected from two-dimensional coordinate value [of the right focus in four frames already extracted in the dynamic-image train using this property] (x1 and y1) - (x4 and y4) is verified.

[0045] It asks for linear combination multiplier ai, bi, (i=1, and 4) with a least square method from the two-dimensional coordinate of the focus to the image of four frames, and a new coordinate (X, Y). The error Res which calculates (x', y') for an approximation coordinate from the coupling coefficient obtained by reverse (X, Y) is calculated shortly.

[0046]

[Equation 1]

Error Res shows the coordinate error of the new focus. Since the error has arisen in the extract location when an error is larger than a threshold, this data is not effective as study data.

[0047] Above (i) - (iii) It is also possible to judge by whether close distribution of the average luminance of a face field or brightness is in the criteria range besides a valuation basis.

[0048] The study data judged that study image storage section 16b is the above-mentioned validity are stored. If the stored number of study data reaches the number of regular, a study image will be sent to principal-component-analysis section 16c.

[0049] In principal-component-analysis section 16c, principal-component-analysis (KL expansion) application is carried out to the image data stored in study image storage section 16b, and it asks for a proper image. The proper image of a high order N individual is stored in the dictionary image 17 from the one where characteristic value

is larger.

[0050] Although the above-mentioned example explained the example which performs dictionary generation on real time from a dynamic image, it is also possible to carry out, once it stores in memory only the frame by which the detecting signal from a predetermined actuation detecting element came from the dynamic image. In this case, since possibility effective in dictionary generation is aimed only at a high frame, improvement in the speed and stability of processing can be measured.

[0051] Moreover, although the subspace method which is the technique like a pattern is applied to discernment in the above-mentioned example, you may transpose to structural technique.

[0052] (Explanation of actuation of this equipment 10) Next, actuation of this example is explained along with drawing 7 .

[0053] The first access performs the usual log in which enters a password etc. Or it is also possible to perform an automatic log in using the dictionary of the user who registered in advance.

[0054] In the face field extract section 12, a face field is extracted from the image first inputted from the image input section 11, and it is sent to the situation-recognition section 13. Moreover, in the predetermined actuation detecting element 18, predetermined actuation detecting signals, such as mouse actuation, keyboard grabbing, and button grabbing, are sent to the situation-recognition section. In the situation-recognition section 13, the change in recognition mode and dictionary generation mode is performed based on the time existence of a face field. Moreover, it controls whether study data are collected based on a predetermined actuation detecting signal.

[0055] time of day T0 to T1 sets it as recognition mode first -- having -- the last user -- it identifies whether you are a principal.

[0056] (i) A recognition mode face field is sent to face focus extract section 14a, and a pupil, a nostril, and **** are detected. Next, normalization is performed on the basis of these focus. Similarity with a user's dictionary image is calculated last time by pattern-matching section 14c, and when similarity is higher than a reference value, a normalization image is identified the principal who was using it last time, and becomes accessible. Access is forbidden when smaller than a reference value.

[0057] Next, from time of day T1, it is set as dictionary generation mode.

[0058] (ii) When the dictionary generation mode predetermined actuation detecting element 18 detects predetermined actuation and the predetermined actuation detecting signal is turned on, study data are collected and a dictionary image is generated. The

face field image extracted from the image inputted continuously is sent to face focus extract section 14a, and a pupil, a nostril, and **** are detected. The extracted focus information is sent to frame assessment section 16a, for example, an error is calculated from a formula (3). When an error is smaller than a threshold, it is judged that it is effective as study data, and it is stored in study image storage section 16b. If the study data more than convention number of sheets are stored in study image storage section 16b, study data will be sent to principal-component-analysis section 16c, and a dictionary image will be generated.

[0059] If a user ****, a screen lock will be started immediately. If a user takes a seat again, it will return to the first processing.

[0060] In addition, in the above-mentioned explanation, when the predetermined actuation detecting element 18 detected predetermined actuation and the predetermined actuation detecting signal was turned on, the dictionary image was generated, but further, after a predetermined actuation detecting signal suits at OFF, a dictionary image may be continuously generated between fixed periods (for example, for 1 minute).

[0061] (Example of modification) Next, the example in service of not only the security of an information terminal but an information terminal is explained.

[0062] For example, when registering a face, it considers memorizing simultaneously not only the characteristic quantity obtained from a face field but the actuation hysteresis of an information terminal. while operating discernment of the human being, simultaneously a terminal the information on a face, simultaneously by registering, it can extract from hysteresis what kind of information was searched for.

[0063] If information retrieval is made into an example, a user will do the activity which inputs a retrieval type etc. In addition to information required for identification, the hysteresis of the retrieval type and the content of retrieval is stored in the database. Based on the stored data, the user's interesting content is presumed and extracted and it holds by performing new information retrieval.

[0064] And again, when a user performs a log in to an information terminal, and an actuation return, the service of him which offers automatically information considered to be the request newly searched to the terminal is attained.

[0065] By linking such hysteresis information with the registration information on a face, and being stored, it may become instead of the time stump when to have retrieved such information, it may recognize whether it is close to the condition of the face at the event of when based on this, and the information based on the hysteresis information on of those days may be shown.

[0066] Moreover, as a verification device in the case of registration of a face, if expression detection and recognition are used, it is also possible to take the actuation hysteresis of the information according to human being's feeling condition, and the dictionary of common always, the dictionary when laughing, the dictionary when being angry, etc. can also be generated. In case face recognition is newly performed, information offer corresponding to expression information may be performed.

[0067] When these perform continuously renewal of face registration, and renewal of information actuation hysteresis simultaneously, clearer information offer is attained.

[0068] Although these examples explained the information terminal to the example, they may be transposed, for example to home electronics, such as TV and a telephone.

[0069] In the case of TV, a subminiature camera is carried in remote control. Predetermined actuation is defined as the actuation which is pushing the carbon button of remote control. The actuation which in the case of a telephone micro CCD carries in an earphone and raises an earphone is defined as predetermined actuation.

[0070] Furthermore, the equipment 10 of this example may be used for an automatic teller's machine (henceforth ATM).

[0071] namely, the principal, a user,, by using this equipment 10 for a check, it forgets be based on others and the unauthorized use of usurpation of cash etc. can be prevented certainly. In this case, the image input section 11 is arranged near the control panel of ATM, and it is made for a user's face to be reflected certainly. Moreover, the account number and updating registration time of day may be registered into a face image and coincidence.

[0072]

[Effect of the Invention] As mentioned above, a new face image is extracted out of the image which was inputted according to this invention, in order to register only the input face image which judged whether it was the face image with which this deserves dictionary registration, and was judged to be effective, registration of an unnecessary face image can be eliminated and dictionary precision can be raised.

[Brief Description of the Drawings]

[Drawing 1] It is the block diagram of the equipment of this example.

[Drawing 2] It is drawing showing a time change in principal identification mode and dictionary generation mode.

[Drawing 3] It is the block diagram of the recognition section.

[Drawing 4] It is drawing showing the example of the normalization processing on the basis of a pupil and a nostril.

[Drawing 5] It is the block diagram of the dictionary generation section.
[Drawing 6] It is the explanatory view which used the error at the time of carrying out linear combination.
[Drawing 7] It is drawing explaining actuation of this example.
[Description of Notations]
11 Image Input Section
12 Face Field Extract Section
13 Situation-Recognition Section
14 Recognition Section
15 Dictionary Generation Section
16 Access-Control Section
17 Dictionary Image
18 Predetermined Actuation Detecting Element

[Claim 1] In the face image registration equipment which has a dictionary means by which a person's face image is registered A face field extract means to extract the input face image which is an image of said person's face field out of the image inputted by image input means to input said person's image, and said image input means, A decision means to judge whether the input face image extracted by said face field extract means is a face image which can be registered into said dictionary means, Face image registration equipment characterized by consisting of a registration means made to register into said dictionary means with said decision means by using as said face image the input face image judged that registration is possible.
[Claim 2] Said decision means is face image registration equipment according to claim 1 characterized by extracting the characteristic quantity of a face from the input face image extracted by said face field extract means, comparing this extracted characteristic quantity with the criteria set up beforehand, and judging that registration is possible if said characteristic quantity is said beyond criteria.
[Claim 3] It is face image-registration equipment according to claim 1 characterized by to judge whether it has a detection means of operation detect predetermined actuation of said person, and said decision means can register as said face image for the input face image extracted by said face field extract means between fixed periods after the detection while said detection means of operation has detected said predetermined actuation to register with said dictionary means.
[Claim 4] The image input step which inputs said person's image in the face image registration approach for registering a person's face image, The face field extract step

which extracts the input face image which is an image of said person's face field out of the image inputted in said image input step, The face image registration approach characterized by the input face image extracted in said face field extract step consisting of a decision step which judges whether it is the face image in which said registration is possible, and a registration step into which the input face image judged that registration is possible in said decision step is made to register as said face image.

[Claim 5] Said decision step is the face image registration approach according to claim 4 characterized by extracting the characteristic quantity of a face from the input face image extracted in said face field extract step, comparing this extracted characteristic quantity with the criteria set up beforehand, and judging that registration is possible if said characteristic quantity is said beyond criteria.

[Claim 6] It is the face image-registration approach according to claim 4 characterized by to judge whether it has the detection step of operation which detects predetermined actuation of said person, and the input face image extracted in said face field extract step between the fixed period after detection can register said decision step as said face image while having detected said predetermined actuation in said detection step of operation.

Fig.1

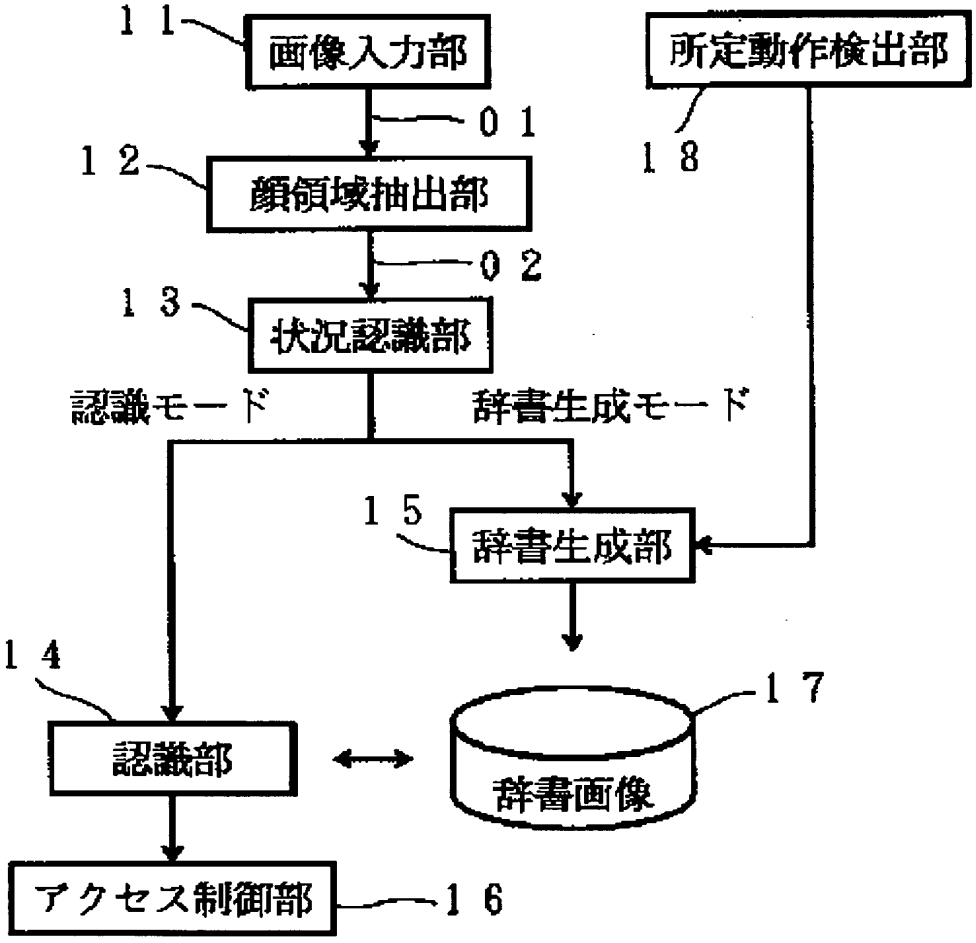

11
画像入力部
所定動作検出部
01
18
12
顔領域抽出部
02
13
状況認識部
認識モード
辞書生成モード
15
辞書生成部
14
17
認識部
辞書画像
アクセス制御部
16

Fig.2

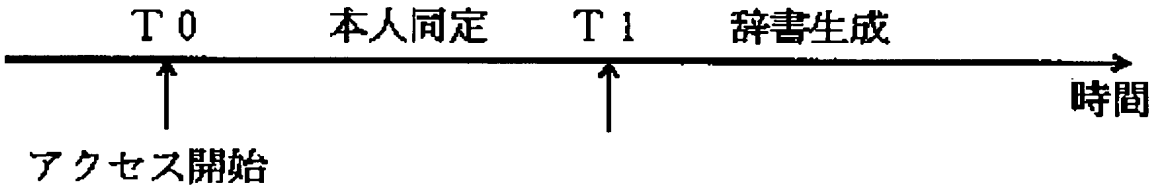

Ｔ０
本人同定
Ｔ１
辞書生成
時間
アクセス開始

Fig.3

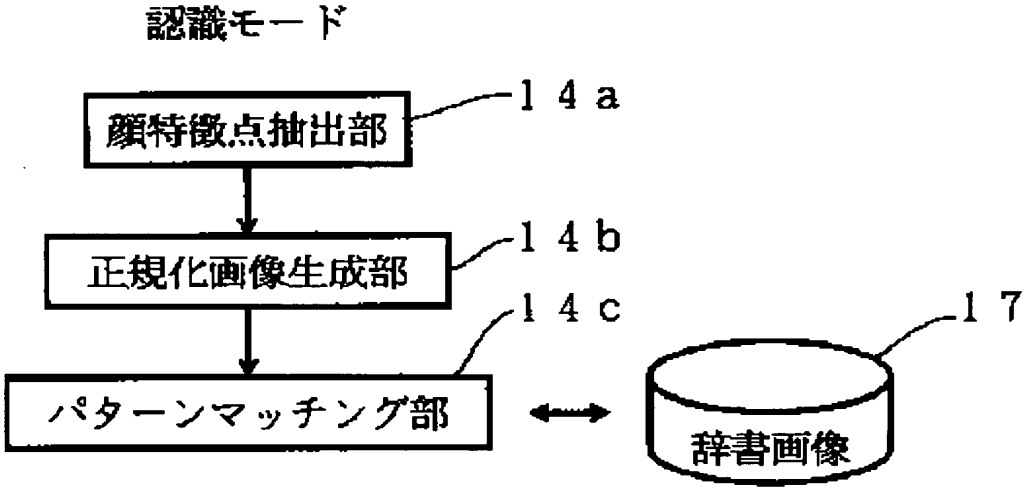
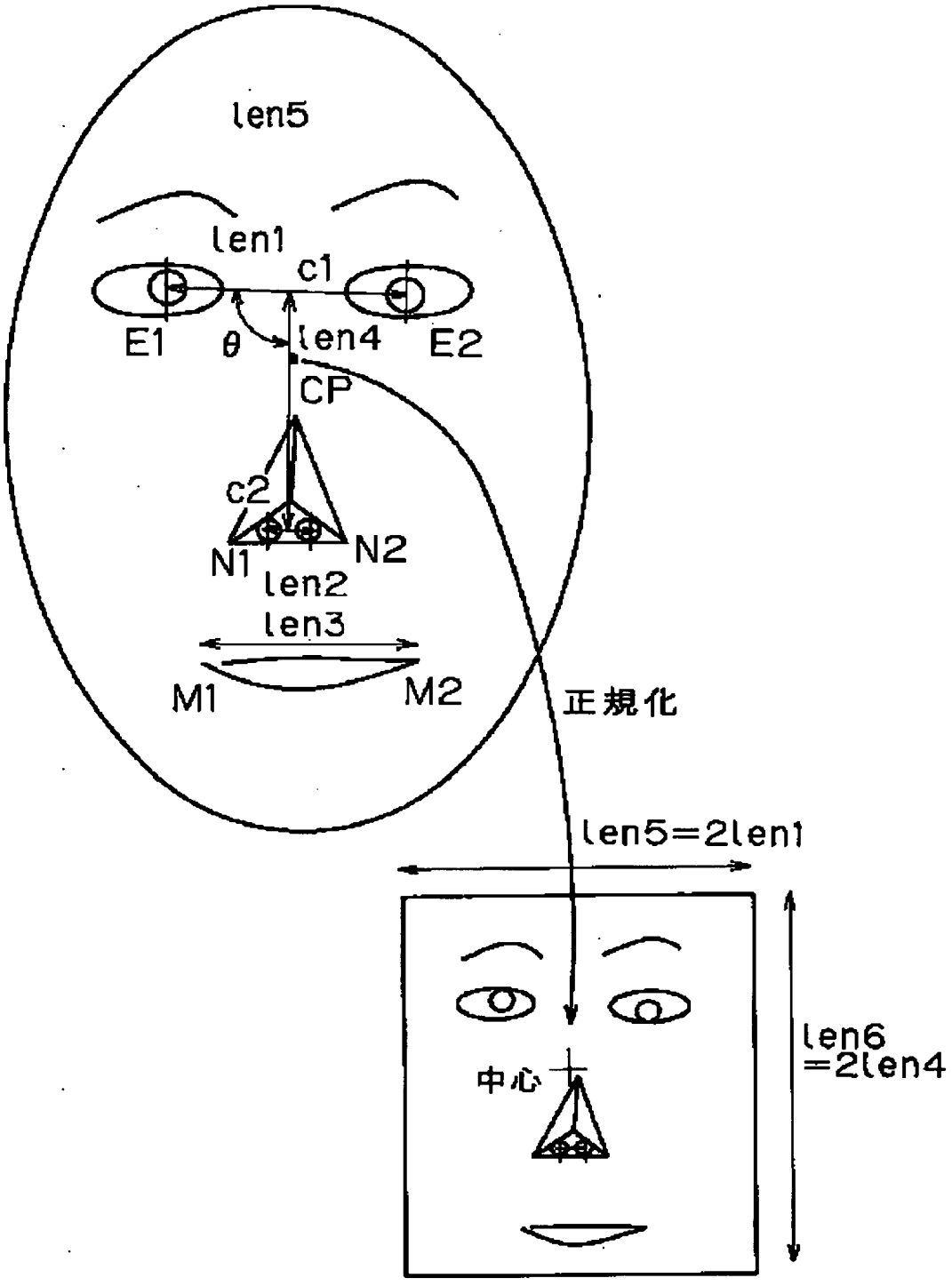

認識モード
顔特徴点抽出部
１４ａ
正規化画像生成部
１４ｂ
１４ｃ
パターンマッチング部
１７
辞書画像

Fig.4


len5
len1
c1
E1
θ
len4
E2
CP
c2
N1
N2
len2
len3
M1
M2
正規化
len5=2len1
中心
len6
=2len4

正規化画像

Fig.5

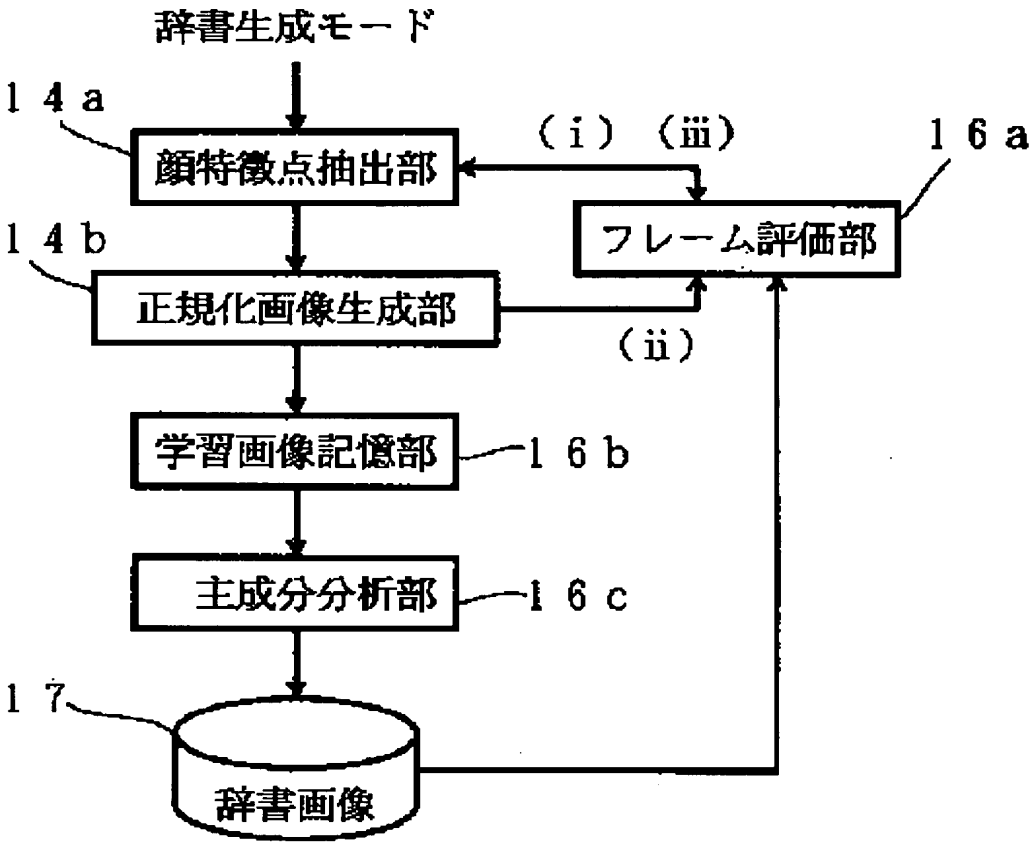
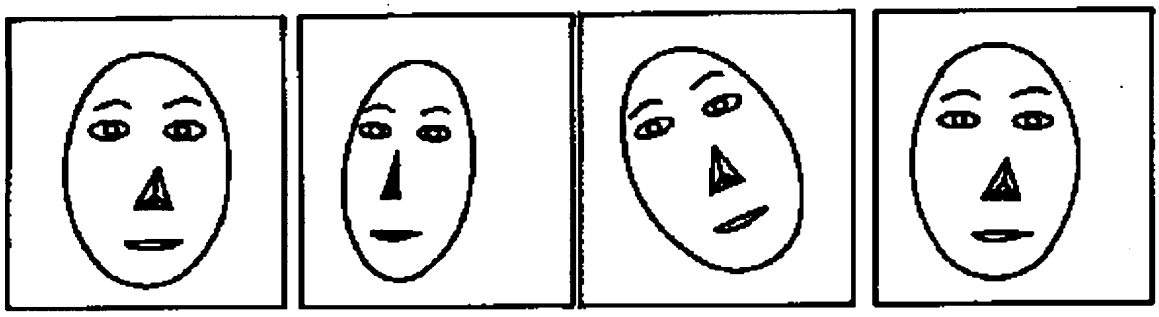

辞書生成モード
14a
顔特徴点抽出部
(i) (iii)
16a
フレーム評価部
14b
正規化画像生成部
(ii)
学習画像記憶部
16b
主成分分析部
16c
17
辞書画像

Fig.6



基底画像１　基底画像２　基底画像３　基底画像４

線形和

☆ 特徴点

任意の方向の顔画像

特徴点の座標（x、y）

$$X=a1\cdot x1+a2\cdot x2+a3\cdot x3+a4\cdot x4$$
$$Y=b1\cdot y1+b2\cdot y2+b3\cdot y3+b4\cdot y4$$

４点以上で係数 $\begin{matrix}a1-a4\\b1-b4\end{matrix}$ を決定